Short Paper
Int. J. on Recent Trends in Engineering and Technology, Vol. 8, No. 1, Jan 2013

Reliability Demonstration Testing for Continuous-Type Software Products Based on Variation Distance

Kiyoshi Sawada
Department of Policy Studies,
University of Marketing and Distribution Sciences, Kobe, Japan
Email: Kiyoshi_Sawada@red.umds.ac.jp

***Abstract*—Reliability demonstration testing for software products is performed for the purpose of examining whether the specified reliability is realized in the software after the development process is completed. This study proposes a model of reliability demonstration testing for continuous-type software such as that for controlling production lines and operating systems of computers. Testing time and acceptance number of software failures are designed based on variation distance. This model has less parameters to be prespecified than the statistical model.**

***Index Terms*—software reliability, reliability demonstration testing, continuous-type software, variation distance**

I. Introduction

The studies of software reliability evaluation can be classified into those of analyzing data obtained in the development process of a software product and those based on software reliability demonstration tests after the development process. As to the former, there are studies on estimating the reliability evaluation measure in the development process or estimating development periods needed to achieve the intended reliability [1, 2]. On the other hand, when the development process is completed, a reliability demonstration testing which has been developed for hardware products originally [3] should also be applied to software for the purpose of examining whether the specified reliability has been attained [4].

We have suggested models for reliability demonstration testing of continuous-type software such as that for controlling production lines and operating systems of computers [5]. Testing time and acceptance number of software failures are designed based on the concept of a statistical test, which requires us to specify the values of producer's and consumer's risks. This study proposes a model of reliability demonstration testing for continuous-type software on the basis of variation distance [6], which can be regarded as a measure to express the distance between two probability distributions. This model can design the test more easily than the statistical model since this model includes less parameters to be prespecified than the statistical model.

II. Notation and Assumptions

The reliability of continuous-type software is discussed in terms of MTTSF (Mean Time To Software Failure), MTBSF (Mean Time Between Software Failures), SFR (Software Failure Rate), and the number of remaining faults in the software. The continuous model for software reliability demonstration testing deals with such a type of software. Under the continuous models, the software of interest is tested for time t, and is accepted if the number of software failures in the test does not exceed an integer s and rejected otherwise. In this case, the design variables are t ($t > 0$) and s ($s = 0, 1, 2, ...$).

The notation used in this paper is as follows:

θ_0 MTBSF on the contract

θ_1 Tolerable lower limit for MTBSF ($\theta_1 < \theta_0$)

E_1 Event that the number of software failures during the test exceeds s

$\overline{E}_1$ Complement of E_1

The assumptions made throughout this paper are listed below:

(i) No fault is removed during the test. All the faults for software failures during the test are removed after the test is completed.

(ii) The software failure times follow an exponential distribution with mean θ, i.e. the number of software failures on (0, t] follows a Poisson distribution with mean t/θ.

(iii) The values of θ_0 and θ_1 are specified at the beginning of the test.

Assumption (ii) is relatively conventional in the software reliability models.

III. Statistical Model

In this section, we describe a model based on the concept of a statistical test with a view to determining the values of t and s in the above SRDT.

In the above SRDT, the probability that event $\overline{E}_1$ occurs when $\theta = \theta_0$ is is given by

$$\Pr[E_1 \mid \theta_0] = 1 - \sum_{i=0}^{s} \frac{(t/\theta_0)^i}{i!} e^{-\frac{t}{\theta_0}}. \tag{1}$$

Likewise, the probability that event $\overline{E}_1$ occurs when $\theta = \theta_0$ is expressed by

$$\Pr[\overline{E}_1 \mid \theta_1] = \sum_{i=0}^{s} \frac{(t/\theta_1)^i}{i!} e^{-\frac{t}{\theta_1}}. \tag{2}$$

The probability in Eq. (1) is called a producer's risk in SRDT as well as in sampling theory, while that in Eq. (2) is

DOI: 01.IJRTET.8.1.60

called a consumer's risk. The producer's and consumer's risks, respectively, signify the probabilities of Type I and Type II errors in terms of a statistical test.

When the values of the producer's and consumer's risks are specified to be equal to or less than α and β, respectively, a feasible region for designing a software reliability demonstration test is written as

$$1-\sum_{i=0}^{s}\frac{(t/\theta_0)^i}{i!}e^{-\frac{t}{\theta_0}} \le \alpha, \tag{3}$$

$$\sum_{i=0}^{s}\frac{(t/\theta_1)^i}{i!}e^{-\frac{t}{\theta_1}} \le \beta. \tag{4}$$

In general, the simultaneous equations obtained by replacing inequalities in Eqs. (3) and (4) by equalities do not always have a unique solution since s is an integer. A practical solution would be obtained by using the following three conditions:

(i) The producer's risk is strictly equal to α.

(ii) The consumer's risk does not exceed β.

(iii) Acceptance number s is the minimum.

IV. Variation Distance Model

Let F_1 and F_2 denote two types of discrete probability distribution, and p_{1i} and p_{2i} (i = 0, 1, 2, ...) be probability mass functions associated with F_1 and F_2, respectively. Then the variation distance [6] of F_1 and F_2 is defined by

$$d_v(F_1, F_2) \equiv \sum_{i=0}^{\infty} | p_{1i} - p_{2i} |. \tag{5}$$

The variation distance in Eq. (5) can be regarded as a measure to express the distance between F_1 and F_2. In other words, as the variation distance increases, we can distinguish F_1 from F_2 more easily.

This section presents a model based on the variation distance for the purpose of determining the values of t and s of the continuous model.

When $\theta = \theta_0$, we have

$$\Pr[E_1 | \theta_0] = 1-\sum_{i=0}^{s}\frac{(t/\theta_0)^i}{i!}e^{-\frac{t}{\theta_0}}, \tag{6}$$

$$\Pr[\overline{E}_1 | \theta_0] = \sum_{i=0}^{s}\frac{(t/\theta_0)^i}{i!}e^{-\frac{t}{\theta_0}}. \tag{7}$$

In the case of $\theta = \theta_1$, we have

$$\Pr[E_1 | \theta_1] = 1-\sum_{i=0}^{s}\frac{(t/\theta_1)^i}{i!}e^{-\frac{t}{\theta_1}}, \tag{8}$$

$$\Pr[\overline{E}_1 | \theta_1] = \sum_{i=0}^{s}\frac{(t/\theta_1)^i}{i!}e^{-\frac{t}{\theta_1}}. \tag{9}$$

A set of Eqs. (7) and (8) represents a probability distribution F_1, expressed by a probability mass function p_{1i} (i = 0, 1) in Eq. (5) when $\theta = \theta_0$. A set of Eqs. (8) and (9) expresses another probability distribution F_2, characterized by a probability mass function p_{2i} (i = 0, 1) in Eq. (5) in the case of $\theta = \theta_1$. Hence, the variation distance of these two probability distributions is written as :

$$d_v(F_1, F_2) = 2\left[\sum_{i=0}^{s}\frac{(t/\theta_0)^i}{i!}e^{-t/\theta_0} - \sum_{i=0}^{s}\frac{(t/\theta_1)^i}{i!}e^{-t/\theta_1}\right]. \tag{10}$$

We can obtain an optimal pair of values, $(t, s)^*$ by maximizing $d_v(F_1, F_2)$ in Eq. (10) since we can distinguish F_1 from F_2 more easily as the variation distance increases.

V. An Optimal Value s^* for Each Value of t

This section shows the existence of the optimal value s^* for each value of t.

Let $D(s) = d_v(F_1, F_2)$, the difference of $D(s)$ is given by

$$D(s+1) - D(s) = 2\left[\frac{(t/\theta_0)^{s+1}}{(s+1)!}e^{-t/\theta_0} - \frac{(t/\theta_1)^{s+1}}{(s+1)!}e^{-t/\theta_1}\right]. \tag{11}$$

The sign of $D(s+1) - D(s)$ follows that

(i) If

$$s < \frac{t(\theta_0 - \theta_1)}{\theta_0\theta_1(\log\theta_0 - \log\theta_1)} - 1, \tag{12}$$

then $D(s+1) - D(s) > 0$.

(ii) If

$$s = \frac{t(\theta_0 - \theta_1)}{\theta_0\theta_1(\log\theta_0 - \log\theta_1)} - 1, \tag{13}$$

then $D(s+1) - D(s) = 0$.

(iii) If

$$s > \frac{t(\theta_0 - \theta_1)}{\theta_0\theta_1(\log\theta_0 - \log\theta_1)} - 1, \tag{14}$$

then $D(s+1) - D(s) < 0$.

Therefore, Let

$$\psi = \frac{t(\theta_0 - \theta_1)}{\theta_0\theta_1(\log\theta_0 - \log\theta_1)} - 1, \tag{15}$$

then we have the following theorem.

Theorem 1:

(i) If ψ is not integer, there exists a unique s^* that maximizes $D(s)$ and s^* is the minimum integer which is greater than ψ.

(ii) If ψ is integer, there exist two s^* that maximizes $D(s)$ and s^* are ψ and $\psi + 1$.

VI. Relation Between Statistical Model and Variation Distance Model

This section reveals relation between the statistical model in Section III and the variation distance model in Section IV.

$d_v(F_1, F_2)$ in Eq. (10) is expressed as following:

$$d_v(F_1, F_2) = 2 - 2\left(\Pr[E_1 | \theta_0] + \Pr[\overline{E}_1 | \theta_1]\right) \tag{16}$$

using the producer's risk $\Pr[E_1 | \theta_0]$ of Eq. (1) and the consumer's risk $\Pr[\overline{E}_1 | \theta_1]$ of Eq. (2).

DOI: 01.IJRTET.8.1.60

Int. J. on Recent Trends in Engineering and Technology, Vol. 8, No. 1, Jan 2013

Therefore, maximizing the variation distance $d_v(F_1, F_2)$ is equivalent to minimizing the sum of the producer's risk $\Pr[E_1 \mid \theta_0]$ and the consumer's risk $\Pr[\overline{E}_1 \mid \theta_1]$ in the statistical model. This means to minimize the expected cost in case the cost of the producer's risk is equal to the cost of the consumer's risk.

VII. Numerical Examples

Fig. 1 shows numerical examples of variation distance $d_v(F_1, F_2)$ when acceptance number of software failures s varies in $s = 0, 1, 2, ..., 10$ for $t = 1000, 2000, 3000, 4000$, and 5000 where $\theta_0 = 2000$ and $\theta_1 = 1000$.

For each value of $t = 1000, 2000, 3000, 4000$, and 5000, the optimal values are $s^* = 0, 1, 2, 2$, and 3, respectively.

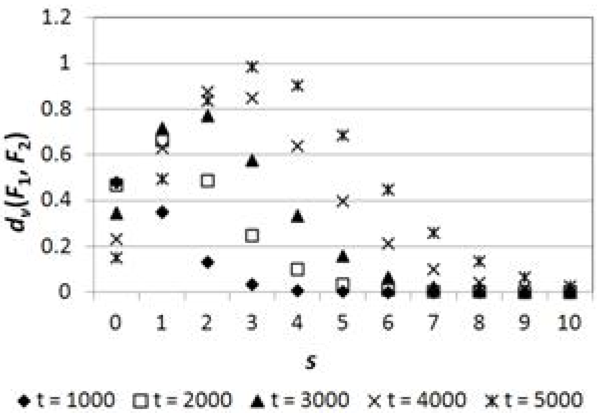

Fig. 1. Numerical examples of variation distance

Conclusions

This study proposed a model of reliability demonstration testing for continuous-type software such as that for controlling production lines and operating systems of computers. When MTBSF on the contract θ_0 and tolerable lower limit for MTBSF θ_1 are given, testing time t and acceptance number of software failures s are designed by maximizing $d_v(F_1, F_2)$. This model has less parameters to be prespecified than the statistical model. Theorem 1 reveals the existence of the optimal value s^* for each value of t.

References

[1] A. L. Goel and F. B. Bastani eds., Special issue on software reliability, IEEE Transactions on Software Engineering, vol.SE-11, pp.1409–1517, 1985.

[2] A. L. Goel and F. B. Bastani eds., Special issue on software reliability: Part II, IEEE Transactions on Software Engineering, vol.SE-12, pp.1–181, 1986.

[3] N. R. Mann, R. E. Schafer, and N. D. Singpurwalla, Methods for Statistical Analysis of Reliability and Life Data, New York: John Wiley & Sons, pp.404–410, 1974.

[4] H. Sandoh, "Reliability demonstration testing for software", IEEE Transactions on Reliability, vol.40, pp.117–119, 1991.

[5] K. Sawada and H. Sandoh, "A summary of software reliability demonstration testing models", International Journal of Reliability, Quality and Safety Engineering, vol.6, pp.65–80, 1999.

[6] S. Amari, et al., Differential Geometry in Statistical Inference, Institute of Mathematical Statistics, 1987.

DOI: 01.IJRTET.8.1.60